Psychometric properties of the Mandarin version of the Childhood Autism Spectrum Test (CAST)

Psychometric properties of the Mandarin version of the Childhood Autism Spectrum Test (CAST): An exploratory study

Autism Spectrum Conditions (ASC) are characterised by impairments in social interaction and communication, alongside unusually repetitive/restricted interests, behaviour and activity (American Psychiatric Association, 2000). *The International Classification of Disease, 10th revision* (ICD-10) (World Health Organisation, 1993), describes ASC as including four subgroups: Childhood Autism, Atypical Autism, Asperger's Syndrome and Pervasive Developmental Disorder, unspecified. Due to the lack of biomarkers, the diagnosis of ASC still depends on behavioural descriptions (Magyar & Pandolfi, 2007). There has been an emerging consensus that ASC are dimensional disorders representing the upper extremes of one or more quantitative traits, and the traits are likely to be continuously distributed in the population (Constantino & Todd, 2003; Mandy & Skuse, 2008; Spiker, Lotspeich, Dimiceli, Myers, & Risch, 2002). Early studies for this proposed dimensional concept investigated the association between genetic liability for ASC and the milder, non-psychopathological features within the relatives of individuals with ASC (Bolton et al., 1994). Further studies suggested that autistic traits are continuously distributed in the larger general population (Constantino & Todd, 2003; Posserud, Lundervold, & Gillberg, 2006). Factor analysis has been used to explore the possible dimensions of ASC, with the data being collected using screening and diagnostic assessment instruments in different samples (Bolte, Westerwald, Holtmann, Freitag, & Poustka, 2011; Happe & Ronald, 2008).

A single autism dimension was proposed based on findings from a factor analysis of the psychometric structure of the Autism Diagnostic Interview-Revised (ADI-R) and the Social Responsiveness Scale (SRS) in both clinical and population samples (Constantino et al., 2004; Constantino & Todd, 2003). Other researchers, also using factor analysis, have suggested that the impairments of ASC might be multidimensional (Lecavalier et al., 2006). Three studies using data from the ADI-R proposed a three-factor solution for the structure of ASC. Although the structures of these proposed factors differed among the studies, these factors roughly covered the three core domains of impairment in ASC including Social Interaction, Communication, Repetitive and stereotyped behaviours and interests (Georgiades et al., 2007; Lecavalier et al., 2006; Van Lang et al., 2006). Another two studies based on the ADI-R suggested that a two-factor solution was sufficient to explain the autism continuum, including both the Social/Communication and Restricted/Stereotyped Behaviours subgroups (Frazier, Youngstrom, Kubu, Sinclair, & Rezai, 2008; Snow, Lecavalier, & Houts, 2009). A recent study using the SRS suggested a two symptom dimensions including the social communication/interaction and restricted/repetitive behaviours in a sample of ASC and non-ASC siblings (Frazier et al., 2012)..

The lack of consistency in the number and structure of factors identified by previous studies could be partly explained by the heterogeneous research designs used (Kuenssberg, McKenzie, & Jones, 2011; Mandy & Skuse, 2008). The results could also be influenced by sample characteristics, the nature of the assessment instruments, and by the subjective interpretations used when naming the identified factors in factor analysis (Kuenssberg et al., 2011; Mandy & Skuse, 2008). Although there has been no agreement on the number or structure of latent traits for ASC, exploring the

dimensional approach can help with evaluating screening and diagnostic instruments for ASC. This process could indicate which possible dimensions of ASC have been reflected by the items on the instrument. Such research has been conducted using several screening instruments for ASC. A summary of factor analytic studies of ASC screening instruments is given in Appendix 1.

Most investigations of the psychometric properties of screening instruments for ASC have been conducted among Western populations. Only two studies have investigated the psychometric properties of screening instruments for ASC in the Chinese population. Both studies were based on samples recruited in Taiwan. The Social Communication Questionnaire (SCQ) has been translated into Traditional Chinese and applied to the Chinese population in Taiwan (Gau et al., 2011). This study applied the Chinese SCQ to 317 parents of participants recruited from clinical settings. It proposed a three-factor solution for the Chinese SCQ involving Social Interaction, Repetitive Behaviours and Communication. The psychometric properties of the Chinese version of the Autism Spectrum Quotient (AQ) were examined in a sample of 4,192 children. This consisted of both clinic-based children with ASC and community-based children showing typical development in Taiwan (Lau et al., 2012). It proposed a 35-item, five-factor solution for the Chinese AQ. These five factors were Social Skills, Mindreading, Patterns, Attention to Detail and Attention Switching. No studies have investigated the psychometric properties of screening instruments for ASC in the Chinese population in mainland China.

The Childhood Autism Spectrum Test (CAST) is a screening instrument developed in the UK for use in detecting potential cases along the whole autism spectrum. As many

children with ASC, especially those with subtle or milder symptoms, are usually not identified before entering primary school, the CAST is designed to be applied to primary school-aged children (4-11 years) (Kamio, 2007; Williams, 2003). The CAST has demonstrated good validity (Sensitivity=100%, Specificity=97%) as a screening instrument in Western populations (note: not all the screen negatives were assessed) (Williams et al., 2005). Using the UK as a validation sample, an exploratory factor analysis of the CAST identified four factors: social behaviour and routines, speech and communication, peer relationships, and imaginative play (Williams, 2003). So far, only one study has adopted the CAST to assess autistic traits which was conducted in the UK and reported moderate to high heritability for autistic traits in the general population (Robinson et al., 2011).

In this study, a combination approach has been adopted to explore the factor structure of instruments with binary responses (Sharp, Goodyer, & Croudace, 2006). This approach draws on the equivalence of Categorical Data Factor analysis (CDFA) and Item Response Theory (Khalid, 2011) (IRT). Following this method, this study conducted an exploration of potential factors in a Mandarin version of the CAST in a sample drawn from the Chinese population in both clinical settings and mainstream schools using an exploratory CDFA with model parameters transformed to IRT model parameters to assist with interpretation.

Method

Participants

The CAST was developed for the general population in the UK, where children with ASC are integrated into mainstream schools. However, in mainland China children with ASC are turned away from mainstream schools (McCabe, 2003). Most children with an existing diagnosis of ASC cannot enter mainstream schools but must enter various intervention centres (Sun, Allison, Auyeung, Baron-Cohen, & Brayne, 2012). The sample in the present study was drawn from two groups: 1) a group based in mainstream schools in Beijing, involving the parents of 737 typically developing children in school years 1 to 4 who did not have an ASC diagnosis; 2) a clinic-based group involving the parents of 50 children with an existing diagnosis of autism before this study. The latter were recruited from the Beijing China Disabled Persons' Federation (BCDPF) and the Elim autism rehabilitation centre in Qingdao. The BCDPF takes charge of people with all kinds of disabilities in Beijing and is a local branch of the China Disabled Persons' Federation (CDPF). The Elim autism rehabilitation centre is a private centre in Qingdao that provides interventions for children from different regions of mainland China aged 3-6 years with ASC. All the students in ordinary schools did not have a diagnosis of ASC before this study. All the cases recruited from clinical settings had an existing diagnosis of ASC made by Chinese clinicians using DSM-IV or ICD-10.

The Mandarin CAST

The CAST is a 37-item parental completion questionnaire, of which 31 items can be scored. Each item scores 0 or 1, with the total score ranging from 0 to 31 and higher scores indicating greater symptoms of ASC (Scott, Baron-Cohen, Bolton, & Brayne, 2002a). Used as a screening instrument, the recommended cut-off is 15 (≥15) for

indicating a child is at risk for ASC (Scott, Baron-Cohen, Bolton, & Brayne, 2002b). The CAST was translated from English to Mandarin by the first author, a native Chinese speaker. It was back-translated by two Chinese-English bilingual speakers who are not involved in autism research. The Mandarin CAST was initially piloted in an opportunity sample with ten Chinese parents who had children aged 5-10 years; they were selected from among the outpatients in the Paediatric Department of the Peking University First Hospital (PUFH). The final version was then back-translated and approved by the UK authors. Each item in the Mandarin CAST requires a binary (yes/no) response. The 6 items that did not contribute to the score (items 3, 4, 12, 22, 26, 33) were removed from the dataset before analysis.

Procedures

Parents of children in both the mainstream schools and clinical settings used were informed of the study's purpose and procedures by their institutions and invited to participate in screening using the Mandarin CAST. They were sent packs comprising an information sheet, a Mandarin CAST, a consent form and an envelope in which to return the questionnaire. Ethical approval for this research was obtained from the Ethics Committees of the research affiliations.

Data analysis

The Mandarin CAST was distributed to all children in the two samples for screening purposes. A number of 34 questionnaires were completely missing as they were not returned back and another 59 questionnaires had missing items. 53 (7.4%) had one or

two missing items and six (0.8%) had three to seven items missing. Item 25 had the most missing data (10 were missing). In total, 694 questionnaires were available for analysis, of which 655 came from the mainstream sample and 39 from the clinical sample. Since only completed CAST questionnaires were used in the analyses, statistical methods were used to examine the differences in the characteristics between children who were included and those excluded. Unpaired t-tests and one-way ANOVA were used to compare means, and Chi-square test was used to examine differences in proportions. Whenever the numbers were small, a Fisher's exact test was used.

Traditional factor analytical methods use observed data which are continuous to explore an underlying continuous latent variable (Bartholornew et al., 2002). Factor analysis provides a linear factor combination or principal components for observed scores from tests based on continuous data (Kline, 2000; Williams, 2003). However, these methods are not appropriate for tests with binary item responses (yes/no), such as the Mandarin CAST, due to the fact that factor analysis assumes item responses are on a continuous metric (Angold, Erkanli, Silberg, Eaves, & Costello, 2002; Sharp et al., 2006). If categorical data are analysed using methods for continuous data in factor analysis, the true factor structure may be distorted in a multi-dimensional analysis and the factor loading may be biased in uni-dimensional models (Muthen, 1989; Sharp et al., 2006). When applying linear models to binary data, the predictions generated may not lie within a plausible range (<0 or >1) (McDonald, 1999; Sharp et al., 2006).

In this study, the latent structure of the Mandarin CAST was examined using exploratory CDFA. The analysis was conducted using MPlus 6.0 (Muthen & Muthen,

2010). Due to the fact that the item responses were binary, the correlation matrix used in the CDFA was a tetrachoric correlation. The factor structure of the Mandarin CAST was estimated by a robust weighted least square estimator (WLSMV) (Muthen & Muthen, 2010), which has been show to perform well in similarly sized samples (Flora & Curran, 2004; Muthen, du Toit, & Spisic, 1997).

The number of potential factors was determined using three approaches: 1) extraction of as many factors with eigenvalues greater than 1 (the Kaiser criterion) (Bandalos, 2009), 2) extraction of as many factors that fall below the `elbow' of a scree plot, and 3) the best fitting model as indicated by standard CDFA fit indices. These indices were the Chi-square Test of Model Fit, the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI). Values in fit indices indicative of a good fitting model were RMSEA≤0.06, CFI≥0.95, TLI≥0.95 and SRMR≤0.08 (Browne & Cudeck, 1992; Hu & Bentler, 1999).The criteria for indices of an adequate fit model were RMSEA ≤0.08, CFI ≥0.90, TLI ≥0.90 and SRMR≤0.10 (Browne & Cudeck, 1992; Hu & Bentler, 1999; Muthen & Muthen, 2009). Oblique Geomin rotation was applied to the factor solution. The Geomin method is recommended when it is expected that factors will be correlated and there may be cross-loading factors (Allison, 2009; Muthen & Muthen, 2010). In this study, factor loadings ≥0.35 were considered salient (Floyd & Widaman, 1995). Once the factor solution was chosen, items that did not load saliently (i.e. <.35) on to any of the factors or cross-loaded onto more than one factor and loaded only modestly on both factors (i.e. <0.5) were removed from further analysis. In addition, factors where only one or two items saliently loaded onto them were also removed since at

least three items were required on each factor to ensure each one was well-measured (DeCoster, 1998).

To enable greater insight into the psychometric performance of the Mandarin CAST was represented graphically using two IRT approaches: the Item Characteristic Curve (ICC) and the Test Information Curve (TIC). The graphs for each factor were presented separately. The ICCs provided trace lines for IRT models that were defined by each item's difficulty parameter (related to the intercept/threshold) and discrimination parameter (related to the factor loading). They also showed the probability of responding positively over the full range of the latent trait (Edelen & Reeve, 2007).

The TIC provides a graphical representation of the precision of the measurement – related to reliability – over the full range of the latent trait (Edelen & Reeve, 2007; Sharp et al., 2006). Since information is equal to the inverse of the squared standard error of the measurement, it is useful for indicating where test has the greatest precision for distinguishing amongst individuals (Lord, 1980).

Results

Participants

The mean age of children was 8.2 years old (SD=1.33, range: 4-11.4). After statistical comparison, no differences were observed between the children included in this analysis and those excluded due to missing data. The characteristics of the two samples are shown in Table 1.

[insert Table 1 about here]

The categorical data factor analysis model

All the 31 items that could be scored on the Mandarin CAST were included in the CDFA. Based on the Kaiser criterion, the eigenvalues suggested a factor solution of up to seven factors, while the scree plot suggested a two or possibly three-factor solution. Thus, in the following analysis, models extracting up to seven factors were considered. Model fit statistics were in the adequate range for the two-factor solution, and naturally improved with each additional factor extracted (Table 2).

[insert Table 2 about here]

The model fit statistics for the two-factor solution met the adequate model fit criteria. The first factor included items concerning Social Interaction and Communication (Factor 1), while the second factor included items concerning Inflexible/Stereotyped Language and Behaviours (Factor 2). Sixteen items loaded onto the first factor and thirteen items loaded onto the second factor. There were two items, Noticing unusual details (Item 6) and Unusual memory (Item 19), that did not have salient loadings on either factor. The correlation between these two factors was moderate (Geomin rotated factor correlations (GFC)=0.402). The three-factor solution proposed a third factor, Attention to detail (Factor 3). The third factor only included the two items (items 6 and 19). In order to ensure that factors are measured well, it is recommended that each factor should have at least four or five indicators (observed items) (Muthen

& Muthen, 2009). Thus, only a two factor solution was considered further. Five items were removed from the analysis: item 6, 14, and 19 did not saliently loaded onto either factor, and item 20 and 23 were observed to cross-load modestly on to both factors. The CDFA was re-run with the remaining 26 items for a two-factor solution. The model was still stable and met the goodness of fit indices criteria (RMSEA=0.035, CFI=0.961, TLI=0.954, SRMR=0.063). The correlation between the two factors was moderate (GFC=0.447). The factor loadings for the 26 items are shown in Table 3.

[insert Table 3 about there]

Graphical presentation of IRT

The ICC and TIC of the IRT model for the two factors are shown in Figure 1. With regard to Factor 1, all 15 item difficulties are roughly at the same level for the social and communication trait and all items are located to the right of the figures. This indicates that they are located towards the more severe end of the continuum. The ICCs of Factor 1 indicates that a child located between 1 and 2 standard deviations above the population mean on the latent trait would have a 50% probability of endorsing the Mandarin CAST items. Children at the mean latent trait value (0) have a low probability of endorsing any Factor 1 item in the Mandarin CAST. The shapes of the ICC curves of 15 items are similar with sharp slopes, indicating a high discriminating power for these items with respect to the social and communication trait. The TIC of Factor 1 shows that measurement precision, and thus reliability, is highest around 1.8 standard deviations above the mean of the latent trait. At this point reliability is high at around 0.89, and is also good for people above the mean of latent

trait (Table 4). This has implications for the interpretation of low and very high scores on this trait.

In terms of Factor 2, all 11 item difficulties are located towards the more severe end of the continuum (the right of the figure). The ICCs of Factor 2 indicate that a child located at around 1 standard deviation above the population mean (0) on this factor would have a 50% probability of endorsing the Mandarin CAST items. The TIC of Factor 2 indicates that measurement precision is highest around 1 standard deviation above the mean of the latent trait. Reliability is 0.83 at this point (Table 4).

[insert Figure 1 & 2 & Table 4 about here]

Discussion

The exploratory CDFA suggested the Mandarin CAST measures two latent autistic traits, Social and Communication, and Inflexible/Stereotyped Language and Behaviours. It can be considered as a useful screening instrument of ASC with good reliability to discriminate individuals with ASC from those without ASC. The third factor suggested by the analysis focusing on the unusually good ability to remember details that children with ASC sometimes have. This ability could be considered one of the more unusual talents that have been observed in children with Asperger's Syndrome (Conson, Salzano, & Grossi, 2011; Glanzman, 2010; James, 2010; Sevik et al., 2010). This potential factor has been proposed by two studies based on data from the AQ (Auyeung et al., 2008; Hoekstra et al., 2008). However, there has been a move towards combining the social and communication deficits into a social-communication factor, and considering repetitive behaviours, interests and activities

(RBIA) as another separate factor (Kuenssberg & McKenzie, 2011). The two factors identified in this study were in line with the DSM-V (Ghaziuddin, 2010). One previous study investigated the factor structure of the ADI-R algorithm and suggested a two-factor solution: Stereotyped Language and RBIA, and Impairments in Social Interaction and Communication (Frazier et al., 2008). Another study on the ADI-R algorithm also identified two factors: Social/Communication and Restricted/Repetitive Behaviours (Snow et al., 2009). A similar factor solution was further proposed by one study based on another screening questionnaire, the Autism Asperger Assessment (Baron-Cohen, Wheelwright, Robinson, & Woodbury-Smith, 2005; Kuenssberg & McKenzie, 2011) and SRS (Frazier et al., 2012). The modest correlation between factors was found which suggested that the two dimensions may largely separable. Thus, this study provided some evidence against the single autism dimension.

However, the CAST need to be used with caution. The reliability of the factors it measures was high only between 1 and 2 standard deviations above the mean of the latent trait. The variation of reliability within a range of the latent autistic trait suggested that it may be difficult to use the current CAST to grading levels of autistic symptoms in the general population. Thus, when using the CAST, it is important to remember that low scores do not necessarily indicate the absence of autistic traits. This is a common issue with screening tools that is accentuated by the dichotomous scaling approach of the instrument. The current scaling approach may have limited the measurement ability of the CAST for prognosis and treatment tracking. Further research is needed to assess whether other ASC screening instruments suffer the same problem.

Limitations

A limitation of this study is the representativeness of the study sample. Both children with existing diagnosis of ASC and children in the general population were recruited. Most of the children with an existing diagnosis (Childhood Autism) were autistic children at the more severe end of the spectrum. This sampling approach was also adopted in a previous study in the UK (Auyeung et al., 2008). The combined sample will not be representative of the general population in mainland China. However, the CAST was developed among the general population in the UK, where children with ASC are integrated into mainstream schools. In mainland China, children with ASC are turned away from mainstream schools (Sun, Allison, Auyeung, Matthews et al., 2012) Thus, the mainstream school population in mainland China may not be comparable to the mainstream school population in the UK. The sampling strategy in this study aimed to produce a more heterogeneous sample with an appropriate proportion of individuals with ASC and typically developing children.

Although most factor analysis studies have been conducted among clinical samples, it has been suggested by a previous factor analysis study using the ADI-R that studies examining measurement factor structures should also be performed in more heterogeneous samples with a greater proportion of individuals who do not meet the diagnostic criteria for the ASC (Frazier et al., 2008). In addition, due to missing values, not all the collected questionnaires were used for this analysis. Another approach using data imputation for those missing values could be conducted.

However, 91% of the questionnaires were fully completed and it is unlikely that the excluded questionnaires would affect the results of the factor structure.

Personal judgments were required in the factor analysis during the reduction of items, which may have influenced the results (Allison et al., 2008). These judgments included choosing the criteria for salient loading, the factor extraction and the criteria for indices of model fit. However, the inclusion criteria choices adopted in this study were consistent with previous studies (Frazier et al., 2008; Kanne et al., 2011; Pandolfi, Magyar, & Dill, 2009). In addition, this study is only an exploratory study. Due to the limited sample size, we only perform the exploratory factor analysis and removed items with poor fit to get a fit model. In the future, a confirmatory factor analysis should be performed to confirm the two-factor structure in larger Chinese sample.

Conclusion

A two-factor solution was proposed which comprised Social and Communication, and Inflexible/Stereotyped Language and Behaviours. The Mandarin CAST measured the two latent traits adequately at the point of discrimination. However, the interpretation of the scores needs to be made with caution since there is variation of reliability when measuring the latent traits. The two factors measured by the Mandarin CAST provided some evidence to support the idea that the symptom manifestations of ASC in Chinese children shared some similarities with Western populations.

Acknowledgements

We are grateful to the families who participated in this study. This study was funded by the Waterloo Foundation and the Peking University First Hospital.

References

Allison, C. (2009). *The quantitative checklist for autism in toddlers (Q-CHAT).* (PhD), University of Cambridge, Cambridge, UK.

Allison, C., Baron-Cohen, S., Wheelwright, S., Charman, T., Richler, J., Pasco, G. (2008). The Q-CHAT (Quantitative CHecklist for Autism in Toddlers): a normally distributed quantitative measure of autistic traits at 18-24 months of age: preliminary report. *J Autism Dev Disord, 38*(8).

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders, DSM-IV-TR*. Washington DC: American Psychiatric Association.

Angold, A., Erkanli, A., Silberg, J., Eaves, L., & Costello, E. J. (2002). Depression scale scores in 8-17-year-olds: effects of age and gender. *J Child Psychol.Psychiatry, 43*(8), 1052-1063.

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The Autism Spectrum Quotient: Children's Version (AQ-Child). *J Autism Dev Disord., 38*(7), 1230-1240.

Bandalos, D. L. B.-K., M.R. (2009). *Four common misconceptions in exploratory factor analysis In Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences.* . New York:: Routledge. .

Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The Adult Asperger Assessment (AAA): a diagnostic method. *J Autism Dev.Disord., 35*(6), 807-819.

Bartholornew, D. J., Steel, F., Moustaki, I., & Galbraith, J. (2002). *The Analysis and Interpretation of Multivariate Data for Social Scientists*: Chapman & Hall/CRC.

Bolte, S., Westerwald, E., Holtmann, M., Freitag, C., & Poustka, F. (2011). Autistic traits and autism spectrum disorders: the clinical validity of two measures presuming a continuum of social communication skills. *J Autism Dev.Disord., 41*(1), 66-72.

Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M. (1994). A case-control family history study of autism. *J Child Psychol.Psychiatry, 35*(5), 877-900.

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research, 21*(2), 230-258.

Conson, M., Salzano, S., & Grossi, D. (2011). Neuropsychological functioning of an Asperger child with exceptional skill in arranging picture stories. *Neurocase, 17*(4), 353-359.

Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *J Child Psychol.Psychiatry, 45*(4), 719-726.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Arch Gen Psychiatry, 60*(5), 524-530.

DeCoster, J. (1998). Overview of Factor Analysis. Retrieved Sixth, August, 2012, from http://www.stat-help.com/notes.html

Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research, 16*, 5-18.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9*(4), 466-491.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*(3), 286-299.

Frazier, T. W., Youngstrom, E. A., Kubu, C. S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the autism diagnostic interview-revised. *J Autism Dev Disord., 38*(3), 474-480.

Frazier, T. W., Youngstrom, E. A., Speer, L., Embacher, R., Law, P., Constantino, J. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry, 51*(1), 28-40 e23.

Gau, S. S. F., Lee, C. M., Lai, M. C., Chiu, Y. N., Huang, Y. F., Kao, J. D. (2011). Psychometric properties of the Chinese version of the Social Communication Questionnaire. *Research in Autism Spectrum Disorders, 5*(2), 809-818.

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., Roberts, W. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *J Am Acad.Child Adolesc.Psychiatry, 46*(2), 188-196.

Ghaziuddin, M. (2010). Should the DSM V drop Asperger syndrome? *J Autism Dev Disord., 40*(9), 1146-1148.

Glanzman, M. (2010). Developing Talents: Careers for Individuals with Asperger Syndrome and High-Functioning Autism, 2nd edition. *Journal of Autism and Developmental Disorders, 40*(2), 266-267.

Happe, F., & Ronald, A. (2008). The 'fractionable autism triad': a review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychol.Rev, 18*(4), 287-304.

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. *J Autism Dev Disord., 38*(8), 1555-1566.

Hu, L. T., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling-A Multidisciplinary Journal, 6*(1), 1-55.

James, I. (2010). Autism and Mathematical Talent. *Mathematical Intelligencer, 32*(1), 56-58.

Kamp-Becker, I., Ghahreman, M., Smidt, J., & Remschmidt, H. (2009). Dimensional structure of the autism phenotype: relations between early development and current presentation. *J Autism Dev Disord., 39*(4), 557-571.

Kanne, S. M., Wang, J., & Christ, S. E. (2011). The Subthreshold Autism Trait Questionnaire (SATQ): Development of a Brief Self-Report Measure of Subthreshold Autism Traits. *J Autism Dev Disord.*

Khalid, M. N. (2011). An overview of statistical approaches for assessing model fit. *The international journal of educational and psychological assessment, 8*(2), 69-87.

Kline, P. (2000). *Handbook of Psychometric Testing*. London: Routledge.

Kuenssberg, R., & McKenzie, K. (2011). Confirmatory factor analysis of the Adult Asperger Assessment: the association of symptom domains within a clinical population. *Res Dev Disabil., 32*(6), 2321-2329.

Kuenssberg, R., McKenzie, K., & Jones, J. (2011). The association between the social and communication elements of autism, and repetitive/restrictive behaviours and activities: a review of the literature. *Res Dev Disabil., 32*(6), 2183-2192.

Lau, W. Y., Gau, S. S., Chiu, Y. N., Wu, Y. Y., Chou, W. J., Liu, S. K. (2012). Psychometric properties of the Chinese version of the Autism Spectrum Quotient (AQ). *Res Dev Disabil, 34*(1), 294-305.

Lecavalier, L. (2005). An evaluation of the Gilliam Autism Rating Scale. *J Autism Dev.Disord., 35*(6), 795-805.

Lecavalier, L., Aman, M. G., Scahill, L., McDougle, C. J., McCracken, J. T., Vitiello, B. (2006). Validity of the autism diagnostic interview-revised. *Am J Ment.Retard., 111*(3), 199-215.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* . Hillsdale, NJ: Lawrence Erlbaum Associates.

Magyar, C. I., & Pandolfi, V. (2007). Factor structure evaluation of the childhood autism rating scale. *J Autism Dev.Disord., 37*(9), 1787-1794.

Mandy, W. P., & Skuse, D. H. (2008). Research review: What is the association between the social-communication element of autism and repetitive interests, behaviours and activities? *J Child Psychol.Psychiatry, 49*(8), 795-808.

McCabe, H. (2003). The beginnings of inclusion in the People's Republic of China. *Research and Practice for Persons with Severe Disabilities, 28*(1), 16-22.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: LEA.

Muthen, B. (1989). Multiple-Group Structural Modeling with Non-Normal Continuous-Variables. *British Journal of Mathematical & Statistical Psychology, 42*, 55-62.

Muthen, B., du Toit, S. H. C., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes.

Muthen, L. K., & Muthen, B. (2009, 2009). MPlus short courses. from http://www.statmodel.com/download/Topic%202-v20%20[Compatibility%20Mode]1.pdf

Muthen, L. K., & Muthen, B. O. (2010). *Mplus Users' guide verson 6*.

Pandolfi, V., Magyar, C. I., & Dill, C. A. (2009). Confirmatory factor analysis of the child behavior checklist 1.5-5 in a sample of children with autism spectrum disorders. *J Autism Dev.Disord., 39*(7), 986-995.

Posserud, B., Lundervold, A. J., Steijnen, M. C., Verhoeven, S., Stormark, K. M., & Gillberg, C. (2008). Factor analysis of the Autism Spectrum Screening Questionnaire. *Autism, 12*(1), 99-112.

Posserud, M. B., Lundervold, A. J., & Gillberg, C. (2006). Autistic features in a total population of 7-9-year-old children assessed by the ASSQ (Autism Spectrum Screening Questionnaire). *J Child Psychol.Psychiatry, 47*(2), 167-175.

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happe, F. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Arch Gen Psychiatry, 68*(11), 1113-1121.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002a). Brief report: prevalence of autism spectrum conditions in children aged 5-11 years in Cambridgeshire, UK. *Autism, 6*(3), 231-237.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002b). The CAST (Childhood Asperger Syndrome Test): preliminary development of a UK screen for mainstream primary-school-age children. *Autism, 6*(1), 9-31.

Sevik, A. E., Kultur, E. C., Demirel, H., Oguz, K. K., Akca, O., Ergun, E. L. (2010). Asperger Syndrome with Highly Exceptional Calendar Memory: A Case Report. *Turk Psikiyatri Dergisi, 21*(3), 249-255.

Sharp, C., Goodyer, I. M., & Croudace, T. J. (2006). The Short Mood and Feelings Questionnaire (SMFQ): a unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7- through 11-year-old children. *J Abnorm.Child Psychol., 34*(3), 379-391.

Snow, A. V., Lecavalier, L., & Houts, C. (2009). The structure of the Autism Diagnostic Interview-Revised: diagnostic and phenotypic implications. *J Child Psychol.Psychiatry, 50*(6), 734-742.

Spiker, D., Lotspeich, L. J., Dimiceli, S., Myers, R. M., & Risch, N. (2002). Behavioral phenotypic variation in autism multiplex families: evidence for a continuous severity gradient. *Am J Med Genet., 114*(2), 129-136.

Stewart, M. E., & Austin, E. J. (2010). The structure of the Autism-Spectrum Quotient (AQ): Evidence from a student sample in Scotland (vol 47, pg 224, 2009). *Personality and Individual Differences, 48*(1), 88-88.

Sun, X., Allison, C., Auyeung, B., Baron-Cohen, S., & Brayne, C. (2012). Healthcare and educational provision for autism in mainland China. (In press). *Research in Developmental Disabilities*.

Sun, X., Allison, C., Auyeung, B., Matthews, F. E., Murray, S., Baron-Cohen, S. (2012). Service provision for autism in mainland China: A service providers' perspective. *Res Dev Disabil, 34*(1), 440-451.

Van Lang, N. D., Boomsma, A., Sytema, S., de Bildt, A. A., Kraijer, D. W., Ketelaars, C. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *J Child Psychol.Psychiatry, 47*(1), 37-44.

Williams, J. (2003). *Screening for autism spectrum disorders.* University of Cambridge.

Williams, J., Scott, F., Stott, C., Allison, C., Bolton, P., Baron-Cohen, S. (2005). The CAST (Childhood Asperger Syndrome Test): test accuracy. *Autism, 9*(1), 45-68.

World Health Organisation. (1993). *The ICD-10 Classification of Mental and Behavioural Disorder: Diagnosis Criteria for Research*. Geneva: World Health Organisation.